DOCUMENT RESUME

ED 206 007 CS 206 567

AUTHOR Yinger, Robert J.; And Others
TITLE Selecting Instructional Activities: A Policy-Capturing Analysis.
INSTITUTION Michigan State Univ., East Lansing. Inst. for Research on Teaching.
SPONS AGENCY National Inst. of Education (ED), Washington, D.C.
REPORT NO IRT-RS-103
PUB DATE Jul 81
CONTRACT 400-76-0073
NOTE 29p.
AVAILABLE FROM The Institute for Research on Teaching, 252 Erickson Hall, Michigan State Univ., East Lansing, MI 48824 ($2.50).

EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS *Cognitive Processes; *Decision Making; Decision Making Skills; Educational Research; Elementary School Teachers; *Evaluation Criteria; Grade 4; Grade 5; *Instructional Materials; *Language Arts; Media Selection; *Research Methodology

ABSTRACT

A policy capturing approach was used in a study that investigated teacher judgment during the selection of instructional activities. This approach uses a simple (usually linear regression) model to reproduce the inferential responses of a particular judge. Central to the approach is the manner in which the judge weighs and combines information in the form of discernable cues or features of the objects being judged. Nineteen fourth and fifth grade teachers were given descriptions of 32 language arts instructional activities reflecting five dimensions that teachers use to judge the quality of teaching materials: (1) amount of student involvement, (2) difficulty for students, (3) integration with other skills or subject matter, (4) demand on teacher time, and (5) fit between stated purpose and instructional process. Individual and composite models showed that integration, difficulty, and demand were the most influential judgment dimensions for the teachers. However, the findings suggest that in a realistically complex judgment task teachers use a wider variety of information than can be represented adequately by a linear regression model. (FL)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

U.S. DEPARTMENT OF EDUCATION
NATIONAL INSTITUTE OF EDUCATION
EDUCATIONAL RESOURCES INFORMATION
CENTER (ERIC)

X This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official NIE position or policy.

ED206007

Research Series No. 103

SELECTING INSTRUCTIONAL ACTIVITIES: A POLICY-CAPTURING ANALYSIS

Robert J. Yinger
Christopher M. Clark, and
Merlyn M. Mondol

Published By

The Institute for Research on Teaching
252 Erickson Hall
Michigan State University
East Lansing, Michigan 48824

July 1981

CS206567

This work is sponsored in part by the Institute for Research on Teaching, College of Education, Michigan State University. The Institute for Research on Teaching is funded primarily by the Program for Teaching and Instruction of the National Institute of Education, United States Department of Education. The opinions expressed in this publication do not necessarily reflect the position, policy, or endorsement of the National Institute of Education. (Contract No. 400-76-0073)

Institute for Research on Teaching

Teachers' thoughts and decisions are the focus of studies currently under way at Michigan State University's Institute for Research on Teaching (IRT). The IRT was founded in April 1976 with a $3.6 million grant from the National Institute of Education. That grant has since been renewed, extending IRT's work through September 1981. Funding is also received from other agencies and foundations. The Institute has major projects investigating teacher decision-making, including studies of reading diagnosis and remediation, classroom management strategies, instruction in the areas of language arts, reading, and mathematics, teacher education, teacher planning, effects of external pressures on teachers' decisions, socio-cultural factors, and teachers' perceptions of student affect. Researchers from many different disciplines cooperate in IRT research. In addition, public school teachers work at IRT as half-time collaborators in research, helping to design and plan studies, collect data, and analyze results. The Institute publishes research reports, conference proceedings, occasional papers, and a free quarterly newsletter for practitioners. For more information or to be placed on the IRT mailing list please write to : The IRT Editor, 252 Erickson, MSU, East Lansing, Michigan 48824.

Co-Directors: Judith E. Lanier and Lee S. Shulman

Associate Directors: Lawrence W. Lezotte and Andrew C. Porter

Editorial Staff:

Lawrence W. Lezotte, coordinator of Communications/Dissemination
Janet Flegg, IRT editor
Pat Nischan, assistant editor

SELECTING INSTRUCTIONAL ACTIVITIES: A POLICY-CAPTURING ANALYSIS

Robert J. Yinger, Christopher M. Clark, and Merlyn M. Mondol[1]

Considerable activity and interest has been stimulated in research on teacher thinking by the report of the Panel on Teaching as Clinical Information Processing of the National Institute of Education Conference on Studies in Teaching (NIE, Note 1). To date, research on teacher thinking has been directed at four topics: teachers' planning, judgment, interactive decision making, and implicit theories or perspectives (Clark & Yinger, 1977; Shavelson, Note 2). This paper reports the results of a study of teachers' judgments while selecting instructional activities.

The most frequently used method of studying and representing judgment processes is policy capturing (Slovic & Lichtenstein, 1971; Shulman & Elstein, 1975). This approach begins with a simple (usually linear) model and attempts to reproduce the inferential responses of a particular judge. Of central interest in this paradigm is how judges weigh and combine information in the form of discernable cues or features of the objects to be judged.

The policy-capturing approach has been used in studies of teacher judgments as they relate to characteristics of effective teachers

---

[1]Robert Yinger is a former IRT senior researcher with the Teacher Planning Project and an assistant professor of education at the University of Cincinnati. Christopher Clark is the coordinator of IRT's Teacher Planning Project and an associate professor of education at MSU. Merlyn Mondol is an associate professor of psychology at Saginaw Valley State College, specializing in learning and cognition. The authors wish to express their appreciation to Susan K. Bowns, who provided valuable assistance in the data collection for this study.

(Anderson, 1977), classroom organization (Borko, Note 3), classroom management (Cone, Note 4), instructional strategies (Russo, Note 5), and instructional content (Floden, Porter, Schmidt, Freeman, & Schwille, 1981). The present study adds to this list teacher judgments pertaining to selection of instructional activities. We have advocated elsewhere (Clark & Yinger, 1977) that a greater number and variety of studies are required about teacher judgment of students, about curriculum materials, and about other important aspects of the classroom environment before such research will be useful in policy and training decisions. The present study adds to that data by investigating teacher judgment in realistically complex situations. By applying policy-capturing methodology to judgment situations like those regularly encountered by elementary school teachers, we are also evaluating the usefulness of this methodology for describing the complexities and subtleties of teachers' mental lives.

This study of teacher judgment is one part in a series investigating teacher judgment during the selection of instructional materials. This series includes a study that identifies factors influencing the selection of instructional activities (Clark, Yinger, & Wildfong, Note 6), a policy-capturing study of teacher judgment (reported here), a process-tracing study of teacher judgment (forthcoming), a feature analysis of preferred instructional activities (reported here), and an analysis of teacher's self-reported judgment processes (forthcoming).

The underlying hypothesis of these studies is that the selection of attractive, appropriate, and effective instructional activities is an important step in teacher planning for instruction (Yinger, Note 7). Furthermore, teacher judgment is an important psychological process in this activity-selection task.

These are the major research questions guiding this study.

1. What factors do teachers take into account in selecting instructional activities?

2. How do teachers differ in the relative emphasis they place on factors influential in selecting instructional activities?

3. To what extent is the policy-capturing approach adequate for representing teacher judgments about complex instructional-activity descriptions?

4. What do teachers' judgments reveal about what constitutes a good learning activity?

Method

Subjects

Nineteen fourth-and fifth-grade teachers (five males and 14 females in their late 20s to middle 50s) from two Michigan school districts volunteered (and were paid) to participate in the study. The average teaching experience was 13 years, with a range from five to 33 years. Twelve teachers taught in self-contained classrooms, while seven taught in team-teaching situations or a combination of team-teaching and departmental arrangements. Seven of the teachers taught in urban settings, 11 in suburban communities, and one in a rural area. Although all the teachers were taking graduate courses, only nine of them were in master's degree programs. The rest were taking graduate courses in order to be permanently certified as teachers.

Materials

Thirty-two one-or two-page descriptions of language arts writing activities were derived from activities selected from a commercially available instructional catalog of language arts activities for upper elementary classrooms (Forte, Frank, & McKenzie, 1973). The activity descriptions were all presented in the same general format: an activity

title, a one-or two-sentence statement of the purpose of the activity, and a listing of the steps involved in planning and conducting the activity.

Each activity description was edited to reflect five important dimensions teachers use to judge the quality of language arts instructional materials (Clark, Yinger, & Wildfong, Note 6). These dimensions, or cues, were (1) student involvement, (2) difficulty for students, (3) integration with other skills or subject matter, (4) demand on teachers, and (5) fit between stated purpose and instructional process.

The 32 descriptions were constructed to represent a full factoral matrix of high and low values for each cue. Four researchers independently rated each activity, negotiating when their ratings disagreed.

We hypothesized that the evaluation and selection of instructional materials is not one, but a series of judgments leading to a final decision to implement or not implement an activity. To investigate this process and to better reflect the complexity of the judgment task, each participant responded on a nine-point continuum to four questions about each activity.

1. How *attractive* is this activity to you?
2. How *appropriate* is this activity as part of a catalog of language arts activities for fourth- and fifth-grade teachers?
3. How likely would you be *to use this activity as it is* in your present classroom?
4. How *effective* do you think this activity would be for your students?

These four questions, in effect, represent our hypothetical model of the activity judgment process. In selecting materials for classroom use, (e.g., while browsing through a collection of instructional activities) a teacher may first be personally attracted to an activity, then assess its appropriateness for the content and purpose (s)he has in mind, and finally evaluate its potential value for his/her own students. By phrasing the

four judgment questions relative to different reference groups (i.e., in reference to the teacher, to fourth- and fifth-grade teachers in general and to the teacher's own class), we hoped that the components of the selection process might be more independently visible in the data.

Procedure

After a set of six warm-up activities and a question and answer session, each teacher was given a loose-leaf booklet containing the 32 activity descriptions, which were arranged in a different random order for each teacher. The teachers proceeded through the booklet of activities, responding to the four judgment questions on the reverse side of each activity description.

Data Analysis

The study proposed to assess the degree to which judgments about instructional materials could be modeled by linear mathematical representations. Simple linear regression equations have been shown to provide, in a number of cases, very good explanations of judgmental responses (e.g., Goldberg, 1968). Therefore, we computed linear regression equations for each judge (teacher). The five cues were treated as independent variables on which the ratings given to each case were regressed. Using programs from the Statistical Package for the Social Sciences, 76 regression equations were computed.

The regression equations produced by the analysis were examined in two ways. First, regression weights were studied to determine which factors were significant predictors of each judge's rating of the instructional activities. Squared multiple correlations were used to assess the extent to which a judge's rating was systematically related to the five activity features varied across the cases. Though judgment policy analysis is most

powerful as an intra-individual method, a second analysis was conducted on the mean weights of the regression models for each of the four judgments to examine any group trends.

Results

Individual Judgment Policies

The analysis of teacher judgment produced four judgment models (regression equations) for each of the 19 teachers. Since space limitations do not permit a presentation of all 76 of these models, one judge was selected for illustrative purposes.

Teacher 2 is a fifth-grade teacher in an urban school with 16 years of experience. The mean rating given by this judge to all activity descriptions was 6.22 with a standard deviation of 2.66. The average squared multiple correlation from the regression analysis of the four judgment models was .40 (adjusted R^2 = .28), which indicates that, as a conservative estimate, a little less than one-third of the variations in Teacher 2's ratings can be accounted for by the activity features studied.

For the portion of this teacher's judgment accounted for by the regression analysis, a fairly consistent weighting pattern is present across the four judgments. In Table 1, one can see that Teacher 2's judgments of the activity descriptions were influenced most by the perceived difficulty of the activity for the students (Cue 2: *difficulty*) and the integration of various language arts skills or other subject matter in the activity (Cue 3: *integration*). Also, for three of the four judgments, the fit of the stated purpose of the activity with the described instructional process (Cue 5: *fit*) was a significant influence. The potential student involvement in an activity and the demand that an activity puts on the teacher did

not reliably influence this teacher's ratings. From this description one might predict that Teacher 2 is likely to search for and use language arts activities that are well organized and constructed, that are not too difficult for the students, and that integrate a variety of language arts skills or other subject matter.

Table 1

Regression Weights for Four Judgments

Determined by the Judgment Policy of Teacher 2

Feature	Judgment 1 Attractiveness	2 Appropriateness	3 Likelihood of use	4 Effectiveness
Cue				
1. Involvement	.13	.00	.44	.16
2. Difficulty	-.81**	-.59*	-1.63***	-1.34***
3. Integration	.63**	.59*	.87**	.66*
4. Demand	.13	.22	.13	-.31
5. Fit	.63**	.47	.75*	.72*

Note. *$P < .10$, **$P < .05$, ***$P < .01$.

Composite Judgment Policies

Teacher 2, with judgment models characterized by moderate consistency and three significant regression weights, was quite different from the rest of the judges in this study. Across all four judgment tasks, 44% of the regression equations had no significant regression weights, and a

majority of the significant regression models had low squared multiple correlations (mean R^2 = .18; range from .03 - .50).

To more closely examine the composite judgment policies, average regression weights were computed for the five features that were manipulated in the activity descriptions (see Table 2). For activity attractiveness judgments, the feature most heavily weighted was *Fit*. This was closely followed by the feature *Demand*. *Difficulty* was, on the average, negatively weighted, that is, activities perceived as being too difficult for the students were judged as being less attractive.

Table 2

Mean Regression Weights for Each Feature Cue

Feature	Judgment Task 1 Attractiveness	2 Appropriateness	3 Likelihood of use	4 Effectiveness
Cue				
1. Involvement	.07 (.40)	.02 (.31)	.11 (.43)	.02 (.40)
2. Difficulty	-.16 (.34)	-.04 (.28)	-.14 (.52)	-.17 (.43)
3. Integration	.11 (.50)	.01 (.35)	.15 (.53)	.05 (.42)
4. Demand	.21 (.42)	.15 (.35)	.11 (.48)	.12 (.38)
5. Fit	.24 (.32)	.21 (.22)	.23 (.35)	.28 (.36)

Note. Numbers in parentheses are standard deviations

The average weightings for the second judgment task, that of appropriateness, followed a similar pattern to that of the first judgment. The highest average weighting was given to *Fit* followed by the weighting given to *Demand*. For this judgment, however, *Difficulty* had a negligible negative weighting.

The composite policies for the third and fourth judgment tasks, those of likelihood of use and effectiveness, followed the pattern of the first two judgments. In both cases, teachers gave *Fit* the highest average weight and *Difficulty* a negative weight.

Characteristic Judgment Policies

The composite judgment policies described in the previous section provide some notion of the tendencies among these teachers for weighting certain activity features more highly than others. The fact that the regression weights in Table 2 are averages, however, obscures and often cancels out individual differences among the teachers. The large extent of these differences is reflected in the large standard deviations indicated in Table 2.

To further analyze similarities and differences among the teachers, the standardized regression coefficients for the 43 significant policy equations ($p < .10$) were mapped onto polar coordinate graphs (Figures 1-4). These graphs allow a figural representation and comparison of the significant judgment models for each judgment task. The heavy solid pentagon represents a weighting of zero for each cue, and positive and negative weightings are represented by points outside and inside of the pentagon, respectively. Simple models (having only one significant regression weight) are indicated by narrow solid lines. Since for these models all other weightings are treated as zero, the three sides not tangent to the

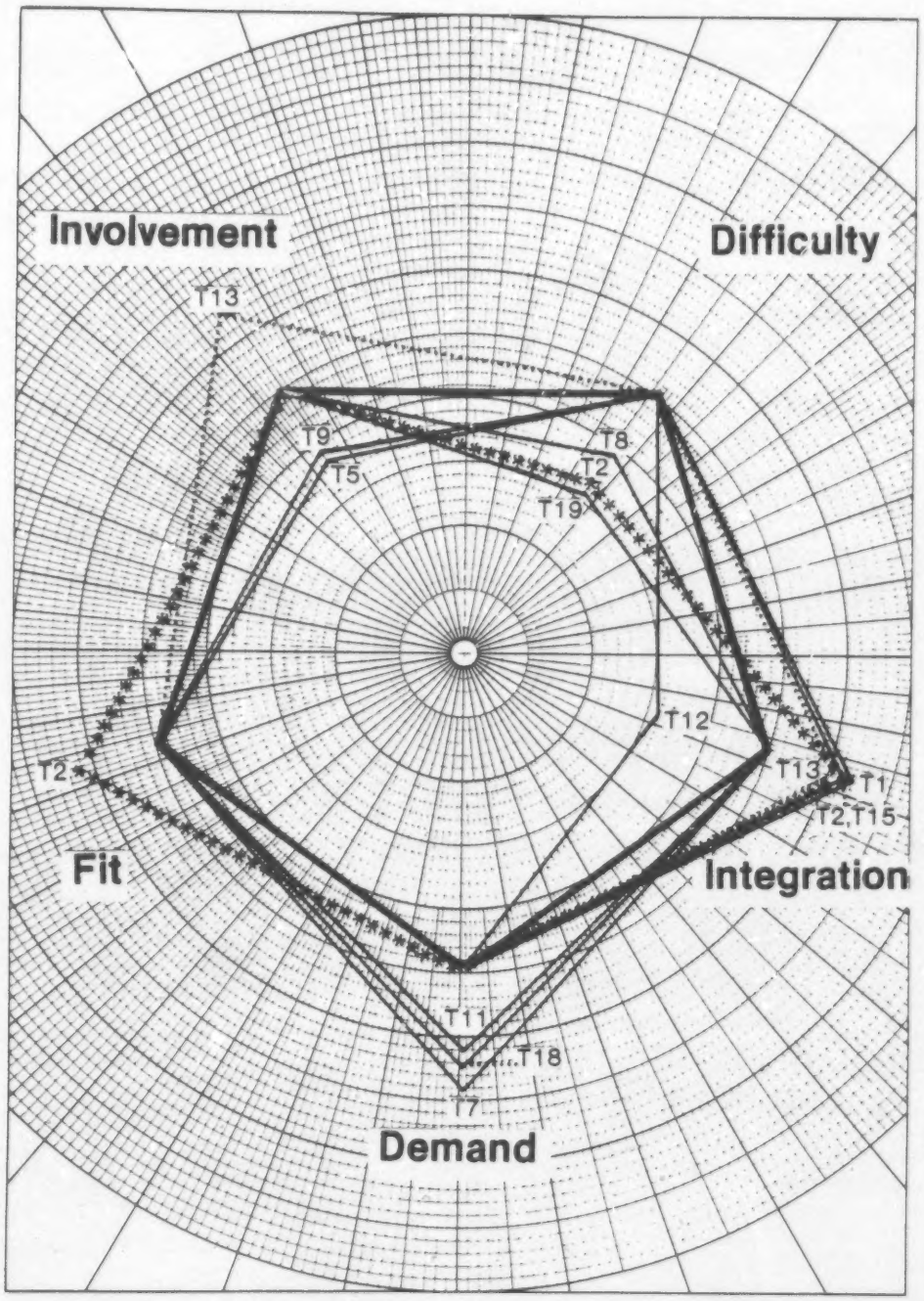

Figure 1. Representation of twelve significant policy equations for Judgement Task 1: Attractiveness.

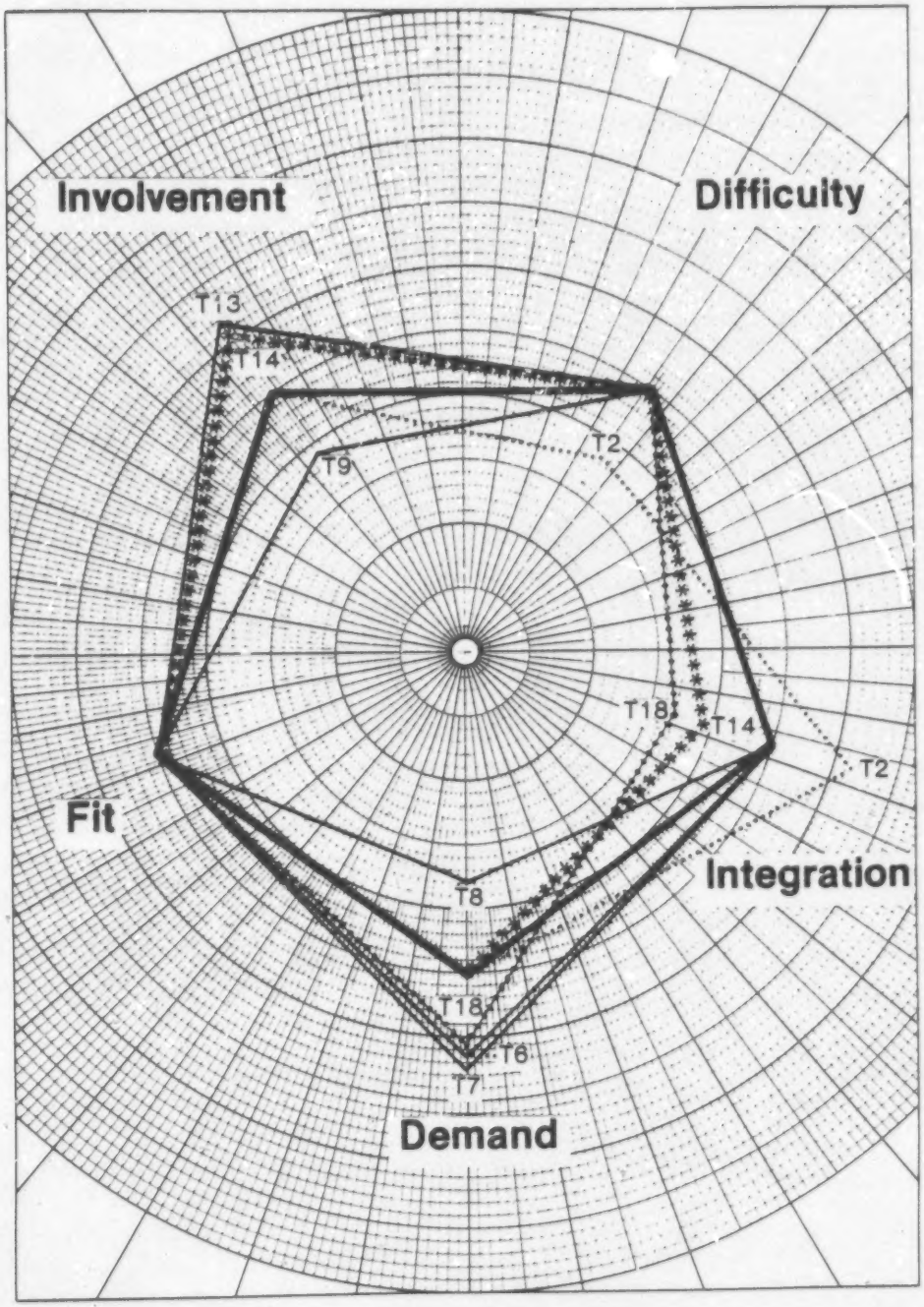

Figure 2. Representation of eight significant policy equations for Judgment Task 2: Appropriateness.

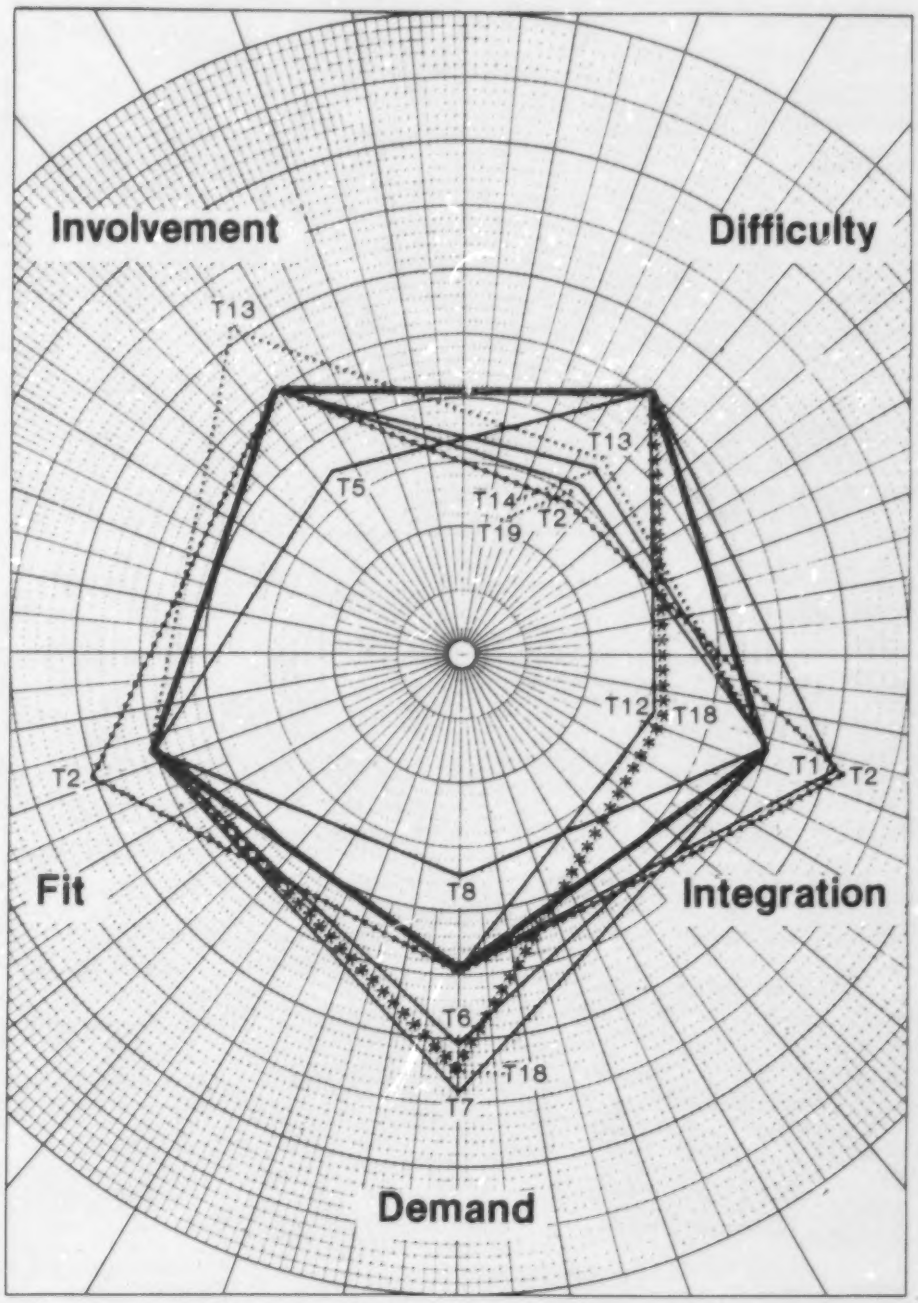

Figure 3. Representation of 11 significant policy equations for Judgment Task 3: Likelihood of use.

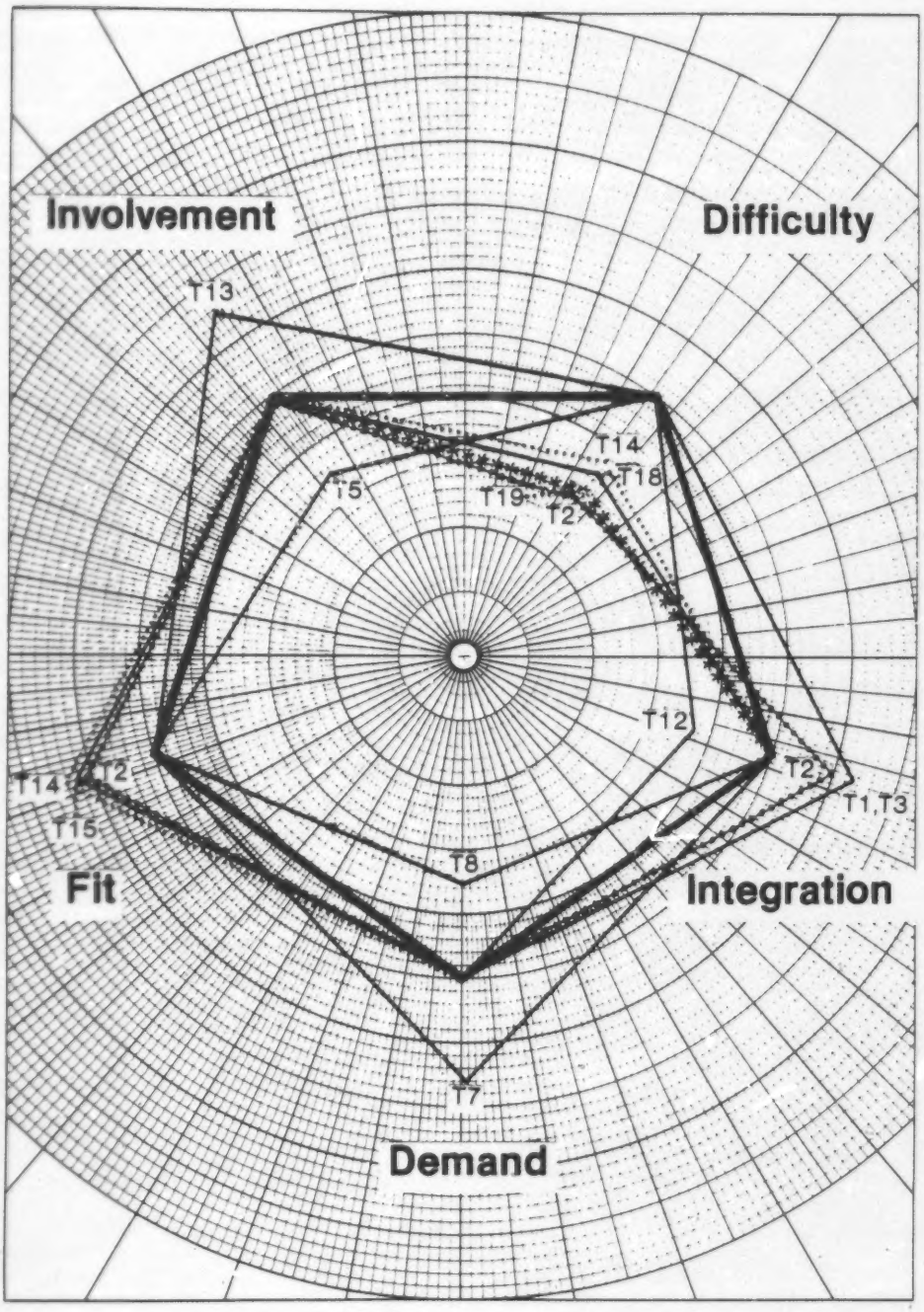

Figure 4. Representation of 12 significant policy equations for Judgment Task 4: Effectiveness.

weighted cue are represented by the heavy solid lines. Complex models (having more than one cue weighted) are indicated in these figues by textured lines.

A variety of information can be obtained from these figures. For instance, Figure 1 represents the 12 significant models for the first judgment task, attractiveness. Of the 12 models, 10 are simple, relying on only one cue. Each of the five cues manipulated in the model has been significantly weighted by the teachers, but the frequency of weightings varies from only one weighting (for *Fit*) to five weightings (for *Integration*). *Fit* and *Demand* were only weighted positively, *Difficulty* was only weighted negatively, and *Involvement* and *Integration* were weighted both positively and negatively. Similar information can be obtained for the other three judgment tasks, and in addition, the point labels allow comparison of the same teacher's judgments across all four judgment tasks. (Only four of the 19 teachers had no significant equation for any judgment task and therefore do not appear on any figure. Two teachers appear on only one figure.)

Since, with the exception of Teacher 2, the models represented in Figures 1-4 have low R^2 values, these graphs may be best used for comparative purposes, indicating the wide degree of individual differences among the teachers. They do, however, suggest some interesting trends in the weighting patterns of the teachers who are represented.

Across the four judgment tasks certain cues were significantly weighted much more often than others. One can see in Table 3, for instance, that *Integration* was significantly weighted 16 times, while *Fit* was significantly weighted only five times. The number of significant weightings for the other three cues fell somewhere in between. It can also be observed that for most of the feature cues, the number of significant regression weights

was fairly consistent across the four judgment tasks. The most obvious deviation from this trend is for *Fit*, where three of the six significant weights assigned to the judgment tasks were given for the fourth judgment task (*Effectiveness*).

Table 3

Significant Regression Weights for the Five Activity Features in Each of the Four Judgment Tasks

Feature	Judgment Task			
	1 Attractiveness	2 Appropriateness	3 Likelihood of Use	4 Effectiveness
Cue				
1. Involvement	3	3	2	2
2. Difficulty	3	1	4	4
3. Integration	5	3	4	4
4. Demand	3	4	4	2
5. Fit	1	0	1	3

Other trends in weighting patterns across the four judgment tasks relate to variations in positive and negative weights assigned to the cues. *Fit* between the stated purpose of an activity and the described instructional process was only assigned positive weightings. This indicates the higher rating of activities having a close fit between its purpose and instructional process.

One-sided weighting was also apparent for *Difficulty*. No teacher significantly weighted this feature positively. This fits expectations that an activity judged as being too difficult would not be rated highly. Note that of the five teachers significantly weighting this cue, only one teacher used it in the second judgment task, appropriateness of including the activity in a set of activities for teachers of fourth or fifth grades.

Involvement, *Demand*, and *Integration*, all had both positive and negative weights assigned to them. One might predict that an activity perceived as having greater student involvment would always be rated more highly than one having low student involvement, but Figures 1-4 show that two of the teachers gave significantly high ratings to activities with low student involvement. One might also predict that activities judged as being demanding on the teacher would not be given high ratings. This is not always true, as Figures 1-4 indicate. Five teachers weighted *Demand* in a significantly positive manner on at least one of the four judgment tasks, and only one teacher weighted *Demand* negatively.

The degree to which other skills or subject matter are integrated into an activity appeared to be an important activity feature for many of these teachers. *Integration* had the greatest number of significant regression weights assigned to it and was weighted negatively as well as positively. One might predict that activities higher on integration would be weighted higher. This was true for the weightings of five teachers. But three teachers assigned negative weight to *Integration* at least once, suggesting a preference for activities that focus on one skill or subject-matter area.

In summary, the simple linear model used here has failed to capture much of the richness and complexity of these teachers' judgments. In those instances where regression equations had any significant descriptive power,

the large individual differences in the models suggest the presence of highly idiosyncratic processes for the selection of instructional activities.

Activity Analysis

Thus far, we have been discussing what this study reveals about the various factors influencing the judgments of these teachers. What do their judgments of the activities suggest about which characteristics of an activity make it especially attractive, appropriate, useful, or effective?

Since the four judgments that the teachers were asked to make about each activity were highly intercorrelated (mean r = .82), we averaged the ratings given by all the teachers for each judgment related to each activity to produce a general indication of quality. The mean rating given to the activities was 6.05, with a range from 3.72 for the lowest-rated activity to 7.55 for the highest rated activity.

Table 4 lists the characteristics of these activities rated highest (in the top 25%) and those rated lowest (in the bottom 25%). Since the 32 activity descriptions were constructed to represent a full factoral matrix of high and low values for each of the five activity features, there is no repetition or replication of any one activity profile. Therefore, at best, the data presented in Table 4 can only suggest the contributing effects of individual activity features.

For the most part, the data in Table 4 reveal no consistent trends for activities in each grouping. The only interesting exceptions appear to be the features *Demand* and *Fit* for the bottom group. All but one of the activities in the bottom group were low in *Fit* and all but two of the activities in this group were low in *Demand*.

In activities of this complexity it is likely that various features counterbalance each other in some configural manner. This seems to be

Table 4

Characteristics of Activities in the Top and Bottom Quartiles of Teacher Ratings

Activity No.	Rank	Activity Features				
		Involvement	Difficulty	Integration	Demand	Fit
5	1	H	L	L	L	L
32	2	L	H	H	L	H
1	3	H	L	L	L	H
12	4	H	L	L	H	H
31	5	L	H	H	H	L
20	6	H	L	H	H	H
16	7	L	L	H	H	L
30	8	L	L	H	L	L
18	25	H	H	H	H	L
6	26	H	L	H	H	L
9	27	H	H	H	L	L
7	28	H	L	H	L	H
23	29	L	H	L	L	L
27	30	L	H	H	L	L
29	31	H	H	L	L	L
25	32	L	L	L	L	L

Note. H = high; L = low

supported by the data in Table 4 so that no feature was unanimously high or low across the two groups. For instance, the activities in the top group were not all high in *Involvement* or high in *Integration*, nor were the activities in the bottom group all low in *Fit*. It is likely that teachers use policies containing tradeoffs like "this activity should be worthwhile even though it may demand a lot from me, because it integrates several important language arts skills."

A difficulty in assessing which features of the activity descriptions contribute most strongly to high or low rankings is related to our method. We chose to control only five features of the activities knowing that there are many more sources of information in activities of this type. For instance, in a previous study (Clark, Yinger, & Wildfong, Note 2) we found that teachers mentioned an average of eight activity features and as many as 11 when making judgments of this type.

Discussion

What can be learned from this study about teachers' selection of instructional activities? The four research questions posed earlier will provide a framework for our answer.

What Factors Do Teachers Take into Account in Selecting Instructional Materials?

Teachers seem to be taking a large number of factors into account in their judgments about instructional materials. The factors or cues that we selected to systematically vary in this study did not fully account for the richness of the teachers' judgments. On the average, less than one-fifth of the variance in teachers' judgments were accounted for by the model used here (although in several cases one-third to one-half of the variance

was accounted for). This is not to say, however, that the feature cues selected for use in this study were not being used by this group of teachers. Fifteen of the 19 participating teachers used one or more of the five cues in a significant manner. For this sample of teachers the factors most frequently influential in their judgments were *Integration*, *Difficulty*, and *Demand*.

How Do Teachers Differ in the Relative Emphasis They Place on Factors Influential in Selecting Instructional Activities?

The wide variations found in cue use are striking. These wide variations made it virually impossible to say anything meaningful about average or group judgment policies. (See for instance the large standard deviations in Table 2.) By focusing on those teachers who had significant policy equations, we were able to get a somewhat better idea about the ways in which they were weighting the five feature cues manipulated in the study. Figures 1-4 provide a much more accurate picture of the variety of individual judgment policies.

Based on these figúes it is also possible to say a few things about trends in cue use. For instance, across the four judgments *Integration* was incorporated most frequently. Also, the weighting for *Integration* was negative as well as positive. Three teachers tended to judge activities more favorably if they did not incorporate other language arts skills, while five different teachers favored more integrated activities. Similar variety was found for *Involvement* and *Demand*. However, when *Fit* was used it was only used positively. *Difficulty* was used only negatively, indicating a lower rating of activities that are perceived as being too difficult.

To What Extent Is the Policy-capturing Approach Adequate for Representing Teacher Judgments about Complex Instructional Activity Descriptions?

Based on the results of this study, policy-capturing methods seem inadequate for modeling a task of this complexity. The low number of significant policy equations and the small amount of variance that could be accounted for attest to some kind of mismatch. One might argue that the fault is not in the method but in how it was used, but we feel that understanding human judgment in contexts as complex as this presents several insurmountable barriers to the use of policy-capturing methods.

The first barrier is the trade-off that one needs to make between task validity (ecological validity) and the limits of the linear model. The linear model functions most effectively in situations in which judgment stimuli are presented in a fully factorial design with cue values represented in an unambiguous manner. As the information presented to a judge becomes richer (in terms of information content) the ability of the linear model to describe the judgment process drops. This decline is largely due to the additional, uncontrolled information available in complex situations that can influence judgments. The trade-off between task complexity and descriptive effectiveness has been shown repeatedly in studies of teacher judgment. For example, in studies of teacher judgment using simplified summaries or profiles of student or curricular characteristics (e.g., Anderson, 1977; Shavelson, Caldwell & Izu, 1977), the linear model has much more effectively captured judgments than in more recent studies that have attempted to use more complex and realistic objects of judgment (e.g., Smith & Glass, Note 8; Byers & Evans, Note 9; Floden et al., 1981). These results, when combined with those of this study, suggest that

policy capturing can only be used effectively in judgment tasks in which the cues can be represented in a simple, unambiguous manner without stripping away the richness and complexity of the real task.

A second barrier is related to the assumptions underlying the linear model. By viewing judgment as a fairly simple additive or multiplicative process where discrete features combine simply with other features, we may be excluding the more holistic and configural components in judgment. It may be that judgment research needs to attempt to accommodate more complex cognitive structures such as schemas, prototypes, and scripts.

A related assumption of the linear model and policy capturing models is that the cue is perceived exactly the way it is presented. Recent research in perception has revealed much about the importance of context and the knowledge and experience that the perceiver brings to the situation. One cannot be sure that what is presented is the same as what is judged. Judgment is as much a perceptual task as it is a conceptual task.

Another limitation inherent in the policy-capturing method is the loss of information about judgments when summarizing across instances. As judgments are averaged, the unique factors influencing each judgment are lost. This is especially true when researchers try to represent the judgments of a group of people.

What Do Teachers' Judgments Tell about What Constitutes a Good Learning Activity?

Analysis of the activities the teachers rated highest and lowest revealed no consistent trends for activities in each grouping. As was previously mentioned, this is most likely due to artificially restricting the number of cues studied. There was a trend for teachers to reject language arts activities that were demanding for the teacher and in which the stated

purposes of the activity did not fit well with the processes prescribed for students. Our comparative analysis of the activities rated in the highest and lowest quartiles indicates that no single activity feature (of the five manipulated) was always present among the highly rated activities or always absent among the activities rated lowest. This suggests that the teachers may have been using a configural process in judging the attractiveness, likelihood of use, and potential effectiveness of language arts activities.

Conclusions

Teacher judgment plays an important part in teacher selection of instructional activities. The policy-capturing analysis presented in this study shows that teacher judgments of the attractiveness, appropriateness, usefulness, and potential effectiveness of language arts activities are highly intercorrelated (i.e., that there are two kinds of activities: those rated high and those rated low) yet there are striking individual differences among teachers in which activities they rate high or low and why. Furthermore, the relatively low descriptive power of a five-cue model of teacher judgment suggests that, in a realistically complex judgment task, teachers use a wider variety of information, possibly in a more configural fashion, than can be adequately represented by a linear regression model.

In an attempt to move beyond the limitations of the linear model in describing complex human judgments, we have conducted a process-tracing study of teacher judgments about the same language arts activities used in the present study. In the process tracing approach, the teacher is asked to think aloud when coming to a judgment. Transcripts of the teacher's deliberations are then analyzed to trace the process whereby (s)he arrived

at a decision. Data analysis is still in progress, but preliminary results suggest that the process-tracing approach provides a more appropriately rich, complex, and configural picture of human judgment in realistically complex situations than does the policy-capturing approach.

This study dealt with one aspect of how curriculum materials, instructional considerations, and human perceptions come together in the minds of teachers as they decide what to teach and how to teach it. We sought to understand why and how these important decisions are made to better understand how the content and process of instruction come to be. Teachers' judgment processes are complex and differ among teachers. We recommend that others use more appropriate, complex methods of studying teacher judgment in action.

Reference Notes

1. National Institute of Education. Teaching as clinical information processing. Report of Panel 6, National Conference on Studies in Teaching, Washington, D.C., 1975.

2. Shavelson, R. J. Research on teachers' pedagogical thoughts, judgments, decisions, and behavior. Unpublished manuscript, Washington, D.C., National Institute of Education, 1980.

3. Borko, H. An examination of some factors contributing to teachers' preinstructional classroom organization and management decisions. Paper presented at the annual meeting of the American Educational Research Association, Toronto, 1978.

4. Cone, R. Teachers' decisions in managing student behavior: A laboratory simulation of interactive decision-making by teachers. Paper presented at the annual meeting of the American Educational Research Association, Toronto, 1978.

5. Russo, N. A. Capturing teachers' decision policies: An investigation of strategies for teaching reading and mathematics. Paper presented at the annual meeting of the American Educational Research Association, Toronto, 1978.

6. Clark, C. M., Yinger, R. J., & Wildfong, S. C. Identifying cues for use in studies of teacher judgment (Research Series No. 23). East Lansing, Michigan: Institute for Research on Teaching, Michigan State University, 1978.

7 Yinger, R. J. A study of teacher planning: Description and theory development using ethnographic and information processing methods. Unpublished doctoral dissertation, College of Education, Michigan State University, 1977.

8. Smith, M. L., & Glass, G. V Factors influencing the evaluation of educational programs: Analysis of judgmental policies. Unpublished manuscript, Boulder, CO: University of Colorado, 1977.

9. Byers, J. L., & Evans, T. E. Using a lens-model analysis to identify factors in teacher judgment (Research Series No. 73). East Lansing, Michigan: Institute for Research on Teaching, Michigan State University, 1980.

References

Anderson, B. L. Differences in teachers' judgment policies for varying numbers of verbal and numerical cues. Organizational Behavior and Human Performances, 1977, 19, 68-88.

Clark, C. M., & Yinger, R. J. Research on teacher thinking. Curriculum Inquiry, 1977, 7(4), 279-304.

Floden, R. E., Porter, A. C. Schmidt, W. H., Freeman, D. J., & Schwille, J. R. Responses to curriculum pressures: A policy-capturing study of teacher decisions about content. Journal of Educational Psychology, 1981, 73(2), 129-141. (Also available as Research Series No. 74, East Lansing, Michigan: Institute for Research on Teaching, Michigan State University, 1980.)

Forte, I., Frank, I., & MacKenzie, J. Kids stuff: Reading and language experiences: Intermediate-Jr. High. Nashville, Tenn.: Incentive Publications, 1973.

Goldberg, L. R. Simple models or simple processes? Some research on clinical judgments. American Psychologist, 1968, 23, 483-498.

Shavelson, R. J., Cadwell, J., & Izu, T. Teachers' sensitivity to the reliability of information in making pedagogical decisions. American Educational Research Journal, 1977, 14, 83-97.

Shulman, L. S., & Elstein, A. S. Studies of problem solving, judgment, and decision making. In F. N. Kerlinger (Ed.), Review of Research in Education (Vol. 3). Itaska, Illinois: F. E. Peacock & Co., 1975.

Slovic, P., & Lichtenstein, S. Comparison of Bayesian and regression approaches to the study of information processing in judgment. Organizational Behavior and Human Performance, 1971, 6, 649-744.

Yinger, R. J. A study of teacher planning. The Elementary School Journal, 1980, 80(3), 107-127. (Also available as Research Series No. 18, East Lansing, Michigan: Institute for Research on Teaching, Michigan State University, 1978.